

THE INFLUENCE OF AUDITORY, VISUAL AND  
AUDIOVISUAL MODALITIES IN THE  
INTERPRETATION OF COUGH REFLEX

---

A thesis submitted in partial fulfilment of the  
requirements for the Degree of  
Master of Speech and Language Therapy  
in the University of Canterbury  
by How Hui Teng

---

University of Canterbury

2012

## Contents

List of Figures .....	4
List of Tables .....	5
Acknowledgements.....	6
Abstract .....	7
Chapter 1: Introduction .....	8
Chapter 2: Literature Review .....	10
2.1 Cough and its Pathways .....	10
2.2 Cough Reflex Testing (CRT).....	13
2.3 Cough Response in CRT.....	16
2.4 Reliability and Validity in CRT .....	22
2.5 Acoustic Information to Assist in Decision Making.....	25
2.6 Visual Information to Assist in Decision Making .....	26
2.7 Multimodality Perception in Assisting with Decision Making.....	27
2.8 Aim of Study.....	30
Chapter 3: Methodology .....	31
3.1 Participants.....	31
3.2 Materials .....	31
3.3 Procedure .....	32
3.4 Consensus Ratings by Expert Team .....	33
3.5 Data Analysis .....	35
Chapter 4: Results .....	36

4.1	Demographic Data .....	36
4.2	Inter-rater Agreement.....	40
4.3	Comparison of Participants' Ratings with Consensus Ratings.....	44
4.4	Cues That Guide Clinicians toward Decision Making .....	48
Chapter 5: Discussion .....		52
	How reliable are clinicians in rating cough response in CRT?.....	52
	Does training in CRT impact reliability? .....	54
	Which modality provides the most salient information? .....	54
	What information can be gleaned from the auditory and visual perception that may cue clinicians towards judging a cough response? .....	57
	What is the level of association between individual ratings by SLTs and consensus ratings by the expert team? .....	59
Limitations .....		61
Conclusion .....		62
References.....		63
Appendix.....		72
	Appendix 1. Questionnaire .....	72
	Appendix 2. A view of the Online Site.....	75

## **List of Figures**

<i>Figure 1.</i> A general overview of the cough pathway.....	12
<i>Figure 2.</i> Various methods of quantifying coughing: a) explosive cough sounds; b) cough seconds; c) cough breaths; and d) cough epochs. Numbers represent cough count. ....	20
<i>Figure 3.</i> Instructions provided to participants in each session before ratings begin.....	34
<i>Figure 4.</i> Definitions provided to participants in each session before ratings begin.....	34
<i>Figure 5.</i> Country in which participants were currently working. ....	37

## **List of Tables**

Table 1. Methods of quantification of cough sound and its description .....	19
Table 2. Number of participants that are performing CRT and trained in CRT .....	39
Table 3. Interpretation of strength of agreement based on Kappa values.....	41
Table 4. Overall inter-rater agreement of cough response in auditory only, visual only and audiovisual modality .....	42
Table 5. Inter-rater agreement of cough response as observed through level of training.....	43
Table 6. Comparison of ratings of presence of cough by participants with consensus ratings	46
Table 7. Comparison of ratings of strength of cough by participants with consensus ratings	47

## **Acknowledgements**

I wish to thank the many individuals who have contributed towards this thesis. I would first like to express my utmost gratitude to my supervisor, Dr. Maggie-Lee Huckabee, for her valuable guidance and support in producing this thesis. Without her constant encouragement and astute advice, this thesis would not be possible. I am also very grateful to Prof. Michael Robb for his insightful suggestions and constructive feedback throughout my thesis writing period.

I am indebted to Anna Miles for the video clips and Anishwar Nair for developing the online site for my research work. His user friendly site made it simple for participants to be part of my research.

I am also thankful to the Ministry of Health, Malaysia for offering this wonderful opportunity to pursue my master's degree. My special thanks to my parents, sister, family members and friends for their unending love and moral support.

## **Abstract**

Cough reflex testing (CRT) is used to assess the sensory and motor components of a reflexive cough. When used as an adjunct to the clinical swallowing evaluation, it has the potential to identify individuals who are more likely to aspirate silently in the event of aspiration. It is unknown how reliable clinicians are at interpreting cough responses and the factors that influence this reliability. Therefore, this study ascertained the reliability of Speech Language Therapists in interpreting cough responses in CRT and determined how sensory perception and training influence reliability. Additionally, the study determined a relative measure of CRT interpretation validity through a comparison with 'expert' consensus ratings. A total of 111 clinicians completed an online survey consisting of ratings of 30 audio-, visual- and audiovisual clips over three sessions, and a final questionnaire. Reliability was substantial ( $\kappa=.76$ ) in rating for presence of cough but only fair ( $\kappa=.25$ ) in rating for strength of cough. Clinicians used a variety of auditory and visual cues to assist in decision making. Training did not result in higher reliability. Ratings of individual clinicians were significantly associated with consensus ratings for presence of cough ( $p < 0.001$ ) and strength of cough ( $p < 0.05$ ). In summary, clinicians could reliably and accurately judge presence of cough response, but were poorer at rating the strength of cough.

## **Chapter 1: Introduction**

Aspiration occurs when foreign materials such as food particles, fluids and saliva are inhaled into the larynx and lower respiratory tract (Marik, 2001, p. 2). In healthy individuals, the host defence system reacts to an aspiration event by expelling these materials from the airway through cough and mucociliary action (Langmore et al., 1998; Marik, 2001). However, in patients with oropharyngeal dysphagia, their host defence systems may be compromised, thus leading to failure to clear foreign materials from the airway and raising the risk for aspiration pneumonia (Marik, 2001).

The term 'silent aspiration' has been used to denote the phenomenon in which individuals are unable to elicit spontaneous coughing when aspiration occurs and show no other signs of distress (Ramsey, Smithard, & Kalra, 2005; Smith Hammond & Goldstein, 2006). Silent aspiration may be a result of a sensory impairment or/and neuromuscular control impairment (Ramsey, et al., 2005). It has been estimated that 15-39% of stroke patients with dysphagia aspirate silently (Ramsey, et al., 2005) and 55% of all patients who aspirate are said to be silent aspirators (Garon, Sierzant, & Ormiston, 2009).

Patients who silently aspirate are at a risk of developing aspiration pneumonia, particularly when aspiration occurs frequently and on large amounts of pathogenic material. The absence of a reflexive cough has been identified as one of the three primary risk factors for the development of aspiration pneumonia in stroke patients with dysphagia (Masiero, Pierobon, Previato, & Gomiero, 2008). Pikus et al. (2003) found silent aspirators to have a 13 times greater risk of developing pneumonia when compared to patients who did not aspirate. It would therefore be of value to develop assessments that would accurately and directly assess cough reflex, and indirectly assessing risk for silent aspiration.

Clinicians have used a variety of assessments to assess aspiration. Videofluoroscopic Swallowing Study (VFSS) and Fiberoptic Endoscopic Evaluation of Swallowing (FEES) are



considered the most accurate methods to detect silent aspiration (Hiorns & Ryan, 2006; Ramsey, et al., 2005). However, these two are not without limitations. The main concern with VFSS is the exposure to radiation (Hiorns & Ryan, 2006; McCullough, Wertz, & Rosenbek, 2001) and both VFSS and FEES are invasive. Both of these methods are also expensive and many institutions lack the equipment required for these studies (Wakasugi et al., 2008). Clinical Swallowing Evaluation (CSE) is another assessment commonly used to evaluate swallowing and aspiration risk, but sensitivity and specificity is low as silent aspiration occurs without outward clinical signs (McCullough, Wertz, & Rosenbek, 2001). CSE alone has a sensitivity of 47% and specificity of 86% in detecting aspiration (England et al., 1998).

Cough reflex testing (CRT) is an assessment that directly assesses the integrity of the cough reflex and therefore identifies individuals who are more likely to aspirate silently in the event of aspiration. CRT when used as an adjunct to CSE may help differentiate patients with intact airway protective mechanism. However, before CRT can be used reliably, further investigations are required to determine proper CRT protocol (Morice et al., 2007; Morice, Kastelik, & Thompson, 2001; Wright, Jackson, Thompson, & Morice, 2010) and to determine reliability and validity for detection of aspiration.

## **Chapter 2: Literature Review**

### **2.1 Cough and its Pathways**

Cough is a defence mechanism of the airway that has two main functions: to protect the lungs against aspiration of foreign material and to assist in the removal of excessive secretions from the trachea, bronchi and lungs (Fontana, 2008, p. 193). It has three characteristic phases- the inspiratory phase, the compressive phase and the expulsive phase- the last of which is accompanied by the characteristic sound of cough (Morice, et al., 2007).

Humans are able to produce cough both voluntarily and reflexively, indicating the existence of a voluntary pathway and reflexive pathway (Figure 1, Lee, Cotterill-Jones, & Eccles, 2002). The reflexive cough may be elicited through stimulation of the larynx, trachea and bronchi (Widdicombe & Fontana, 2006). The mucosa and submucosa in these sites are lined with nociceptors and mechanoreceptors which are sensitive to a wide range of chemical and mechanical irritants (Chang, 2006; Fontana, 2008; Widdicombe & Fontana, 2006).

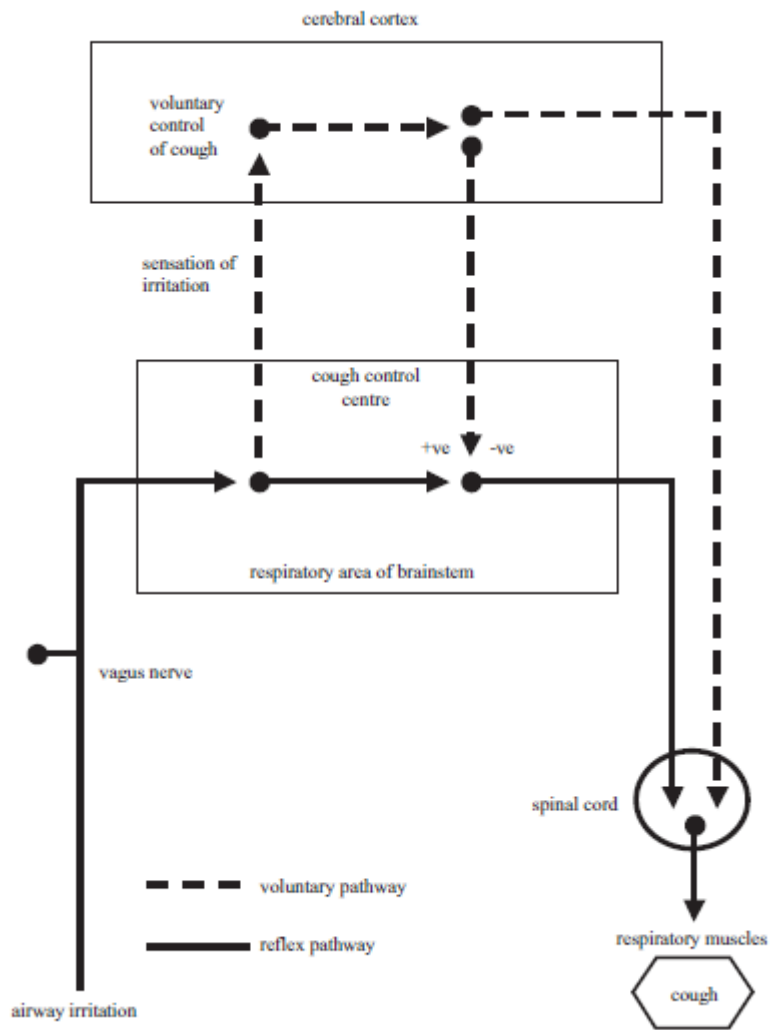
Stimulation of cough receptors by irritants results in generation of action potentials. These signals are sent via the vagus nerve to the second-order neuron at the nucleus tractus solitaries (NTS) located at the brain stem (Chang, 2006; Widdicombe & Fontana, 2006). The NTS then interacts with the cough control centre at the respiratory area of brainstem through its polysynaptic connections, although its exact pathways and connectivity remain unknown (Bolser et al., 2008; Shannon et al., 2004). Signals from the cough control centre are also sent to the cerebral cortex to be cortically modulated though modulation may only occur to a certain extent (Chang, 2006).

It has been suggested that a gating mechanism exists between the NTS and cough network that functions to regulate the behaviour of the cough network (Bolser, et al., 2008).

The gating mechanism functions by permitting inputs from other airway receptors and cortical inputs to interact with inputs from the cough receptors, resulting in increased or reduced excitability of the cough reflex (Widdicombe, 2008). Increased excitability would result in cough threshold being reached and generation of cough (Widdicombe, 2008).

The cough control centre is responsible for producing the cough motor pattern (Shannon, et al., 2004). Signals are carried through the efferent pathways via the spinal cord, recurrent laryngeal nerves and spinal nerves to the various muscles involved in cough execution such as the diaphragm, abdominal muscles, intercostals muscles and the glottis (Chang, 1999; Fontana & Lavorini, 2006). The strength of the stimulus may influence the human response. Weak stimuli may result only in glottal closure while stronger stimuli would result in a cough (Widdicombe & Fontana, 2006).

Humans are also able to voluntarily produce cough which may be in response to a sensation of irritation in the throat (Lee, et al., 2002). As in reflexive cough, sensory information from the cough receptors would be sent to NTS, brainstem and further on to the cerebral cortex. The cerebral cortex may then initiate or suppress the cough response (Lee, et al., 2002). If a cough is to be initiated, the cerebral cortex may send signals via brainstem involvement or directly at the spinal level for activation of the muscles involved in coughing (Lee, et al., 2002).



*Figure 1.* A general overview of the cough pathway. Reprinted from “Voluntary control of cough” by Lee, Cotterill-Jones & Eccles, 2002, Pulmonary Pharmacology & Therapeutics, p. 319.

## **2.2 Cough Reflex Testing (CRT)**

The use of CRT was first reported in the 1950s. The initial purpose of CRT was to assess the effectiveness of antitussive agents and to quantify cough reflex sensitivity (Bickerman, Barach, & Drimmer, 1954). Bickerman et al. (1954) tested different nebulizers and tussigenic agents before determining that citric acid delivered through the Vaponefrin nebulizer provided the most consistent results. Using CRT, normal subjects and patients with bronchial asthma were first assessed for cough reflex sensitivity. This was followed by administration of antitussive drugs and a repeat of CRT. Coughs were monitored and counted for any decrease, which suggests the efficacy of antitussive drugs.

Since then, the CRT has been further developed. Studies have evaluated the most suitable tussive agents (Midgren, Hansson, Karlsson, Simonsson, & Persson, 1992; Pecova, Javorkova, Kudlicka, & Tatar, 2007), delivery device (Wright, et al., 2010), methods of delivery (Morice, et al., 2001) and test end-point (Dicpinigaitis, 2003). A wide variation in CRT procedures existed and the European Respiratory Society (ERS) task force attempted to standardise the CRT by providing guidelines for CRT (Morice, et al., 2007). Some of these guidelines include recommendations for suitable tussive agents to use and test end-points (Morice, et al., 2007). The purpose of CRT has also expanded (Fuller, 2002). It is now used in hypothesis testing (Eckert, Catcheside, Stadler, McDonald, & et al., 2006) and has potential for use in epidemiology studies (Prudon et al., 2005) and clinical practice (Addington, Stephens, Gilliland, & Rodriguez, 1999; Addington, Stephens, & Gilliland, 1999; Wakasugi, et al., 2008).

CRT was first used in clinical swallowing practice by Addington, Stephens, Gilliland, et al. (1999) as a way to identify dysphagic patients who are at risk of developing aspiration pneumonia by assessing the integrity of cough reflex. A total of 161 stroke patients were

recruited for this study. CRT followed by CSE and VFSS were performed on 78 patients while the remaining 83 underwent CRT and CSE but not VFSS. No reason was given on why only 78 of the 161 patients had VFSS. Patients were managed according to CRT results and patients who had an abnormal cough response were put either on restricted diet or non-oral diet. Development of pneumonia was then monitored in all patients though the monitoring time frame was unstated. None of the patients who passed the CRT developed pneumonia, while patients who failed were more likely to develop aspiration pneumonia. The authors also found that CRT had a higher sensitivity and specificity at predicting pneumonia than VFSS. CRT had a sensitivity of 17% and specificity of 100%, while VFSS had sensitivity of 8% and specificity of 95%. As observed, both tests were high in specificity but low in sensitivity.

There were a few other notable limitations in the Addington, Stephens, Gilliland, et al.'s (1999) study. The authors did not have a control group thus no direct comparisons were done between incidence rate of pneumonia in subjects undergoing CRT and in those who did not. The authors also used a protocol that required subjects to be able to follow commands. About 33% of acute stroke patients would present with aphasia (Laska, Hellblom, Murray, Kahan, & Von Arbin, 2001) and may have reduced comprehension thus limiting the usage of CRT in this population. Furthermore, biasing of research results occurs by excluding these patients with severe neurological impairment. A concentration of 20% of tartaric acid was used as the tussigenic agent but no explanation was given on why this concentration was selected. The CRT ended once a cough response was elicited or if the subject did not cough after three inhalations. This might confound the results as subjects who cough on the first inhalation might not reflect a true reflexive cough but as a result of being 'startled' by the presentation (Dicpinigaitis, 2003). Nonetheless, their study gave rise to the idea that

assessment of laryngeal cough reflex through CRT may indicate effectiveness of airway protection.

In a similar study by Addington, Stephens & Gilliland (1999), the authors carried out a prospective study on 604 consecutive acute stroke patients in 2 different rehabilitation hospitals. Four-hundred patients in one hospital were tested for cough reflex using tartaric acid and managed according to the result of the CRT while the other 204 patients from another hospital acted as the control group. The end point of the study was the development of pneumonia. The authors found that patients who were assessed using the CRT and managed according to those test results had a significantly lower likelihood of developing pneumonia. Only 5 of 400 patients who had undergone CRT developed pneumonia compared to the 27 of 204 patients who did not. They concluded that an intact cough reflex indicates a protected airway thus lowering the risk of aspiration and the development of pneumonia. Similar limitations are present in this study as in the prior study by this group: protocols that required subjects to have good comprehension, concentration of tussigenic agent used, response from CRT and end point of study. Although this study had a control group, the subjects for the control group were from a different hospital rendering direct comparisons questionable.

A more recent study was carried out by Wakasugi et al. (2008). The authors investigated the usefulness of the CRT when combined with a modified water swallowing test as a screening tool for patients with silent aspiration. A total of 204 patients suspected of having dysphagia underwent CRT, a modified swallowing test and then either a VFSS or FEES. In the CRT, the patients were asked to inhale citric acid through a face mask, delivered through a nebulizer for one minute. Patients were considered to pass CRT if they coughed at least five times during the test. A diagnosis was made based on the combination of

CRT results and the modified swallowing test results. To determine the validity of CRT in detecting silent aspiration, CRT results were compared with either VFSS or FEES. The authors found that the sensitivity of the cough test for detecting silent aspirators was 0.87, while specificity was 0.89. However, the sensitivity decreased when trace silent aspiration was taken into account (sensitivity of 0.67). When used in combination with modified swallowing test, they were able to accurately predict patients who do not aspirate (89.1%), patients who aspirated but have cough response (73.7%) and silent aspirators (88.2%). Overall, the authors concluded that CRT when used in combination with the modified swallowing test will assist clinicians to distinguish safe-to-swallow patients from patients who are silently aspirating and are unsafe for oral feeding.

Wakasugi et al.'s (2008) study appeared to use a protocol based on previous research but this is not clearly justified in his published research. Some uncertainty about the study includes the concentration of citric acid used. Wakasugi et al. (2008) used only one concentration (1.0w/v %) of citric acid in their CRT and it is unsure why this particular concentration was selected. This concentration does appear to be suitable for detection of larger amount of aspiration but not trace aspiration. It is possible that with a reduced concentration of citric acid, the sensitivity and specificity may change.

### **2.3 Cough Response in CRT**

Cough response may be measured and interpreted in many different ways. Different parameters which may be measured objectively include but are not limited to frequency, intensity, muscular activity and volume changes (Fontana & Widdicombe, 2007). However,



due to factors of time and simplicity, clinicians often use subjective measurement in clinical setting.

One of the approaches that may be used in rating a cough response is through grading of cough. Cough may be classified as eutussia (normal cough), dystussia (pathological), atussia (absent), hypotussia (weakened) and hypertussia (sensitized) (Chung et al., 2009). As the purpose of CRT in clinical practice is to distinguish between patients who have effective defence airway protection with those who do not, a grade of normal, weak and absent may be sufficient (Widdicombe, Addington, Fontana, & Stephens, 2011).

Grading of cough was used in the research by the Addington group (Addington, Stephens, Widdicombe, & Rekab, 2005; Addington, Stephens, Gilliland, et al., 1999; Addington, Stephens, & Gilliland, 1999) but instead of the three grades, the investigators grouped weak and absent together and labelled them as abnormal cough. In their earlier studies (Addington, Stephens, Gilliland, et al., 1999; Addington, Stephens, & Gilliland, 1999), a rather simple definition was provided for these terms. The investigators described a normal cough as “involuntary cough elicited on inhalation” while an abnormal is simply described as “absent or weak involuntary cough on inhalation” (Addington, Stephens, Gilliland, et al., 1999, p. 151).

A more descriptive definition was provided in Addington, et al.’s study in 2005. The authors defined a normal cough reflex as an ‘immediate series of forceful coughs, which are primarily expiratory “airway clearing” in character’ (Addington, et al., 2005, p. 2). A normal cough would indicate intact sensory and motor components of the reflexive cough thus a neurologically protected airway. In contrast, an abnormal cough reflex is defined as “absence of cough, or a diminished (weak) coughing, or coughing not immediately after administration of test stimulus” indicating a neurologically unprotected airway (Addington, et al., 2005, p.

2). These studies provided some general and very subjective guidelines on how to grade the cough (Addington, et al., 2005; Addington, Stephens, Gilliland, et al., 1999; Addington, Stephens, & Gilliland, 1999). It is unknown whether clinicians use any other acoustic or visual information to guide them towards making a judgement.

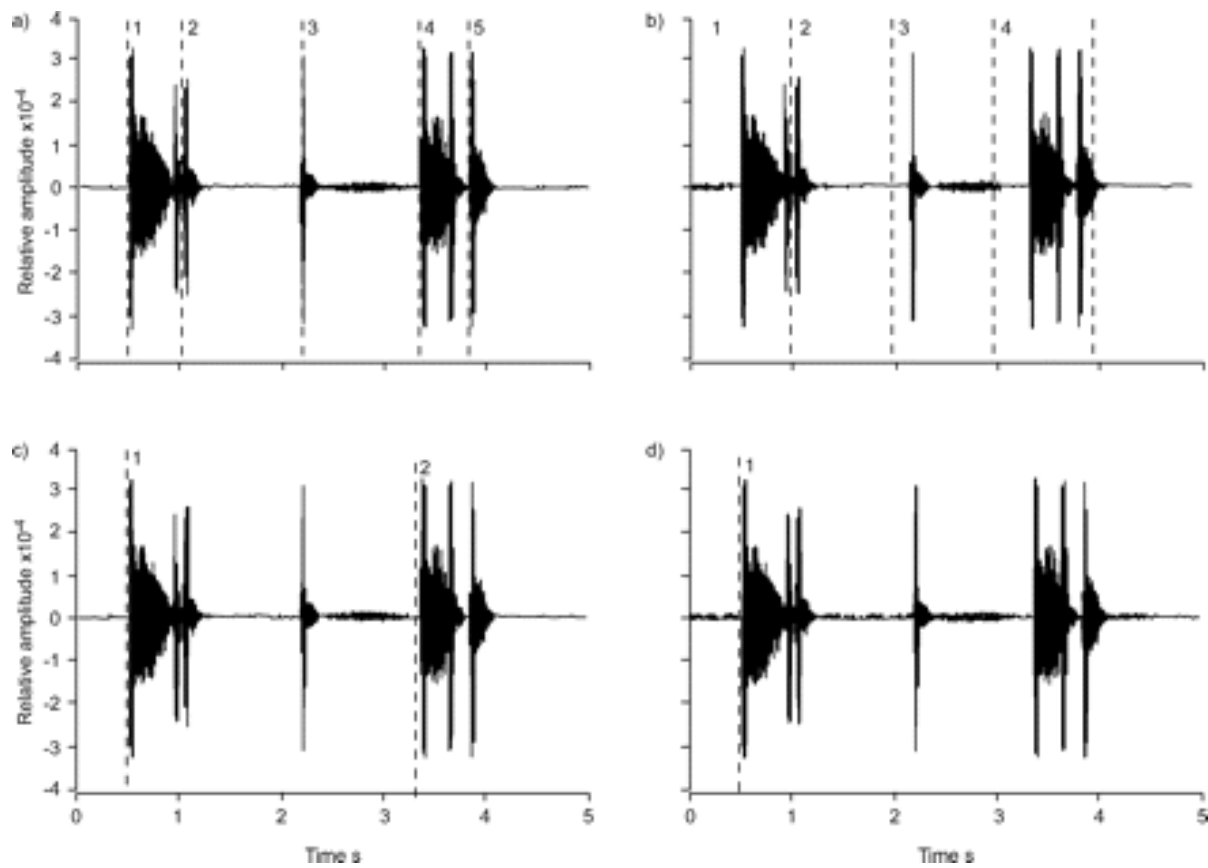
A challenge with grading of cough, especially weak and normal, is that it is done with the assumption that these are dichotomized variables; but in reality, these two grades exist along a continuum. Coughs that are very weak or very strong (normal) would be easier to classify than responses that are moderately weak or moderately strong. It has been found that in dichotomized data with an underlying continuous variable, the chances for misclassification is smaller when values are at the extreme ends of the continuum (Irwig & Groeneveld, 1988). Each clinician may have their own threshold choice in which they judge the grades differently thus opening the possibility for response bias. For example, clinicians who are uncertain of a cough response may grade it as weak instead of normal as the cost of a misclassification may result in a patient acquiring pneumonia after allowing the patient to eat when he/she is unsafe to do so.

Another commonly used subjective measurement of cough in CRT is by quantifying the number of coughs produced. The ERS Task Force (Morice, et al., 2007) have indicated four possible methods on counting cough: through explosive cough sounds, cough seconds, cough breaths and cough epochs. Table 1 provides a more detailed description of the methods and Figure 2 shows examples of quantification based on these four methods.

Table 1

*Methods of quantification of cough sound and its description (as described by Morice et al., 2007 and Kelsall et al. 2008)*

<b>Methods of quantification</b>	<b>Description on how to quantify</b>
Explosive cough sounds	Each characteristic explosive cough sounds is measured.
Cough seconds	The number of seconds containing at least one explosive phase is measured.
Cough breaths	The number of breaths containing at least one explosive cough sound is measured. This method can only be used when breathing is monitored simultaneously.
Cough epochs	The number of periods of continuous coughing, with a 2 s pause or less, is measured.



*Figure 2.* Various methods of quantifying coughing: a) explosive cough sounds; b) cough seconds; c) cough breaths; and d) cough epochs. Numbers represent cough count. Reprinted from “ERS guidelines on the assessment of cough” by Morice et al., 2007, European Respiratory Journal, p. 319

As observed in Figure 2, each method produces a different number of coughs (Morice, et al., 2007). Kelsall, et al. (2008) performed a study comparing three different methods of quantifying cough: counting cough seconds, explosive cough sounds and cough epochs, and found all three methods to be highly correlated. The strongest correlation however was found between explosive cough sounds and cough seconds. Excellent inter and intra-subject agreement for these three methods was also documented. Although all methods may be valid, counting of explosive cough sounds is documented to be the preferred method (Morice, 2008).

Quantification of cough in CRT has been used in clinical swallowing practice to assess for silent aspiration. One such study is by Wakasugi et al. (2008) whereby number of coughs during one minute delivery of citric acid was recorded. More than five coughs were considered to be normal while less than four coughs would be regarded as failing the test. The biggest limitation to this method of measurement is the assumption that all coughs are the same with no difference in intensity or quality of cough. For example, a patient who has weak coughs but is still able to produce five coughs would be considered normal. This may be misleading as the patient may not be able to effectively clear aspirated materials with weak coughs in the event of aspiration. Nonetheless, quantification of cough can be used to determine the presence of a cough response. The ERS task force suggested that either two or five consecutive coughs be taken as a positive response (Morice, et al., 2007). As there are many methods of quantification, it should be made clear on how cough is counted.

## **2.4 Reliability and Validity in CRT**

A variety of assessment tools are used to evaluate the oral-pharyngeal swallowing of suspected dysphagic patients. For a tool to be of any clinical value, it has to be both reliable and valid. Studies have been carried out to assess the reliability and validity of VFSS (Kuhlemeier, Yates, & Palmer, 1998; McCullough et al., 2001), FEES (Colodny, 2002), CBE (McCullough, Wertz, & Rosenbek, 2001) and cervical or bronchial auscultation (Borr, Hielscher-Fastabend, & Lücking, 2007; Leslie, Drinnan, Finn, Ford, & Wilson, 2004).

Limited studies have been performed to determine the validity of CRT in a clinical setting. Addington, Stephens, Gilliland, et al. (1999) assessed the validity of using of CRT in detecting the risk for aspiration pneumonia and found it to be low in sensitivity (17%) but high in specificity (100%). The low sensitivity may be due to the fact that abnormal cough reflex is not the only predictor for aspiration pneumonia. Low functional status, poor oral hygiene, multiple medical diagnosis and smoking coupled with dysphagia have been found to increase the risk for aspiration pneumonia (Langmore, et al., 1998). Accordingly, sensitivity and specificity in detecting for silent aspiration might give a better indication on how well CRT assesses cough reflex integrity. Wakasugi et al. (2008) studied the validity of CRT as a screening tool for silent aspiration with the result of high sensitivity (0.87) and specificity (0.89) to detect for patients who silently aspirated large amount of water. However, the sensitivity is reduced if the purpose is to detect for both silent aspiration in large amounts and trace amounts of water.

Even fewer studies have been performed on the reliability of CRT. Only Addington, Stephens, Gilliland, et al. (1999) reported inter-rater reliability. The authors had six different evaluators evaluate nine subjects though the exact methodological details were not reported. Although they reported inter-rater reliability to be significantly consistent, the authors used

Fisher's exact test, which might not be suitable for analysis of inter-rater reliability. Additionally, sample size used was small (approximately 5% of the subject population).

Decisions based on sound quality are not a novelty in the field of dysphagia. One such example is cervical auscultation. Cervical auscultation is performed in attempts to detect normal and abnormal swallow sounds in patients suspected to have dysphagia and many reliability and validity studies have been completed on this technique. Reliability for identification of normal and abnormal swallow sounds were found to range from poor to fair (Borr, et al., 2007; Leslie, et al., 2004). Clinicians relied on sounds of respiration, swallowing sounds and duration of swallowing sounds to assist them in decision making (Borr, et al., 2007). Levels of experience were not shown to affect the level of reliability (Leslie, et al., 2004) though studies in other areas such as VFSS have shown that training and discussions to agreeable criterion performance improve reliability (Scott, Perry, & Bench, 1998). Studies on validity of cervical auscultation found sensitivity to range from 62% to 94% while specificity ranged from 66% to 90% (Borr, et al., 2007; Leslie, et al., 2004). Practicing speech language therapists (SLTs) were found to be more likely to show a response bias when compared to laypeople or students, by describing a swallowing sound as abnormal when it is in fact normal (Borr, et al., 2007).

There are also reliability and validity studies done outside of the field of dysphagia on judging cough sounds. Chang et al. (2005) completed a study validating subjective ratings of cough quality with bronchoscopy. Parents and clinicians were asked to determine the cough quality in children undergoing elective bronchoscopy. Four cough qualities were described: wet, dry, brassy, non-brassy. Wet cough relates to the presence of secretions while a dry cough is associated with mild or minimal secretions. A 'brassy cough' is usually associated with the presence of tracheomalacia. The authors found that clinicians were able to identify

the cough quality of wet/dry with good sensitivity (0.74) and specificity (0.79) while parents performed marginally worse. However, sensitivity for brassy/ non-brassy cough was lower (0.57) while specificity was good (0.81). Inter-rater clinicians' agreement for wet/dry cough was very good ( $K=0.88$ ) while for brassy/non-brassy cough was good ( $K=0.86$ ). The authors concluded that subjective description of cough has good clinical validity.

Smith, Ashurst, Jack, Woodcock & Earis (2006) evaluated how healthcare professionals describe cough sounds and whether they were able to use this information to accurately make a diagnosis of the underlying disease. Fifty-three participants consisting of physicians and other health professionals were recruited for the study. They were presented with nine spontaneous cough sounds and were asked to 1) identify the subjects' gender 2) select from a list of 10 most common descriptors to appropriately describe the coughs sounds; and 3) make a diagnosis based on the cough sound qualities. To provide an objective comparison, the acoustic properties of cough sounds were analysed to identify coughs with and without mucus and those with and without wheeze. They found participants were able to differentiate cough with or without mucous relatively well ( $76.1\% \pm 14.8$ ) but not cough with wheeze ( $39.3\% \pm 15.0$ ). Participants were also able to correctly identify the gender (93%). Participants were unable to correctly make a diagnosis based on the acoustic quality of cough sound ( $34.0\% \pm 29.0\%$ ).

Based on the combined results of Chang et al. (2005) and Smith, Ashurst, et al. (2006), physicians used the acoustic information from cough to assist in identifying the types of cough and the underlying disease with various successes. Reliability of physicians in identifying wet/dry coughs and brassy/non brassy coughs were similar though sensitivity and specificity for identification of wet/dry coughs were better than brassy/non brassy coughs or wheezy/non wheezy coughs (Chang, et al., 2005; Smith, Ashurst, et al., 2006). However,



information from coughs sounds alone appeared to be insufficient for physicians to make a medical diagnosis. This may be attributed to the fact that different diseases share the same quality of cough sounds. For example, wheezy cough is not only associated with asthma but also chronic obstructive pulmonary disease (Smith, Ashurst, et al., 2006). Laypersons were also able to use the acoustic information to some degree as indicated by their slightly worse results as compared to physician. It was suggested that with training, the ability to identify the characteristics of cough sound may improve (Smith, Ashurst, et al., 2006).

It is uncertain whether results from previous studies on cough and swallow sounds would apply to cough sounds interpretation in CRT. No studies have been carried out to investigate which grading of cough (absent, weak, strong) is easier to identify and how reliable clinicians are in grading coughs. Also, it is of interest to know whether training has an impact on clinicians' reliability to grade coughs.

## **2.5 Acoustic Information to Assist in Decision Making**

Descriptions such as moist, productive, dry and barking have been used by physicians to describe cough sounds before attempting to make a specific medical diagnosis (Chung, et al., 2009). Thus far, no CRT studies have been done looking at the characteristics of cough sounds used to determine the grade of cough. It would be informative to know whether SLTs use certain types of acoustic information to guide them in decision making in CRT. The loudness and force of cough sound may be two of the criteria used by SLTs to decide on a grade. Muscular activity of the abdomen which indicates intensity of cough was found to have a fair correlation with peak decibel level of cough noise (range 0.61-0.84) (Cox et al., 1984). Patients who aspirate have reduced sound pressure level as compared with nonaspirators indicating cough impairment (Smith Hammond et al., 2009; Smith Hammond

et al., 2001). Although reduction in sound pressure level does not directly translate to reduced perception of loudness, it is possible that clinicians would still use the perception of loudness to help in decision making.

## **2.6 Visual Information to Assist in Decision Making**

Little research has been identified regarding how visual input aids in decision making. It is known that video recordings are used to monitor coughing in patients with chronic cough (Smith, Earis, & Woodcock, 2006). The cough sound which coincides with movement associated with cough is used to determine the presence of a cough. However, no description is given on the type of movement associated with cough. Various muscles are activated during the three phases of coughing (inspiratory, compressive and expulsive phase) (Fontana, 2008; Fontana & Lavorini, 2006). For example, during the expulsive phase, the abdominal muscles are activated intensely whereby it is pulled inward to increase the pressure in the airways. This is accompanied by a lowering of the lower ribs (Fontana, 2008). These actions may translate into chest movement and may cue clinicians to the fact that a cough is occurring. In addition to chest movement, the mouth has to remain open during a cough so as to maintain a patent airway (Fontana, 2008). As no research has yet been carried out, it is uncertain how visual information cues clinicians towards the presence and strength of coughs. Additionally, it is uncertain whether visual information could bias clinician in decision making. Patients appearing weak and fragile may bias clinicians' judgement on cough response.

## **2.7 Multimodality Perception in Assisting with Decision Making**

Clinicians evaluating cough response would use information from the auditory and visual modality to assist in decision making. Indeed, humans experience their environment through different sensory modalities and sensory information from all these modalities is integrated to form a consistent and interpretable representation of an event or situation (Davis & Kim, 2004; Doehrmann & Naumer, 2008). It is important to be aware of how these modalities interact with each other as research has shown that information from one sensory modality may affect the perception of another modality (cross-modal interactions) thus modifying the interpretation of an event (Arnold & Oles, 2001; Eramudugolla, Henderson, & Mattingley, 2011; Romei, De Haas, Mok, & Driver, 2011; Schmid & Ziegler, 2006). Cross-modal interactions may either result in cross-modal enhancement which may assist clinicians in decision making or cross-modal illusion which would result in a bias (Alais, Newell, & Mamassian, 2010).

Cross-modal enhancement occurs when the accuracy of perception is increased due to congruent information reaching all modalities (Alais, et al., 2010). In order to have information that is congruent, signals from different modalities have to coincide spatially and temporally (Shams, 2011). In addition, the individual has to have prior knowledge that these signals co-occur in nature (Shams, 2011). Cross-modal enhancement would result in an increased ability to detect an event (Bernstein, Auer Jr, & Takayanagi, 2004; Eramudugolla, et al., 2011), improve comprehension (Arnold & Oles, 2001) and reduce ambiguity (Arnold & Hill, 2001; Arnold & Oles, 2001).

Though no studies have specifically looked at cross-modal enhancement in the perception of cough, research has been done on speech and to a lesser degree, non-speech sounds. Two studies involving speech stimuli were those by Helfer (1997) and Schmid and

Ziegler (2006). Helfer (1997) evaluated the accuracy of speech recognition in noise with stimuli presented either through audiovisual mode or auditory mode alone. Helfer (1997) found improved performance in speech recognition when presentation was through audiovisual presentation. In contrast, Schmid and Ziegler (2006), who studied discrimination of speech and non-speech stimuli through audiovisual mode or auditory mode, found no advantage in the audiovisual modalities compared to auditory modality.

The conflicting results may be due to the differences in methodologies used. One of the differences is the environment in which the studies were conducted. The study by Helfer (1997) incorporated noise in her experiments while Schmid and Ziegler (2006) looked at discrimination tasks with no noise present. According to Callan, Callan, Kroos and Vatikiotis-Bateson (2001), a general rule of multimodality integration is that the greater the degradation of unimodal stimuli, the greater would be the enhancement resulting from a combination of modalities. This partly explains why a cross-modal enhancement is found in the Helfer (1997) study but not by Schmid and Ziegler (2006). Schmid and Ziegler (2006) gave another possible explanation, suggesting that as the error rates in the unimodal auditory task were already low, a ceiling effect prevented further enhancement by audiovisual mode.

The CRT environment may be similar to those found in Helfer (1997) study. Clinical CRT may be done in an environment with background noise in addition to the noise from the nebulizer resulting in a degraded sound signal. An audiovisual condition may then be helpful to the clinician in evaluating the cough. The ceiling effect suggested by Schmid and Ziegler (2006) may only be present when there is no cough and therefore the error in making this judgement would be minimal. In this situation, audiovisual modality would not provide any extra information when compared to auditory modality alone.

It should be noted that the interaction between the auditory and visual modalities may also result in cross-modal illusion. Cross-modal illusion occurs when slight incongruent information reaches different modalities and is partially integrated resulting in an illusion (Shams, 2011). One of the famous studies in cross-modal illusion is by McGurk and McDonald (1976). In this study, the authors presented incongruent audiovisual stimuli and asked participants to repeat what they had heard. An example would be the sound 'ba' paired with a visual recording of a model saying 'ga'. In this particular example, 98% of adult participants responded that they heard the syllable 'da'. This effect has been termed the 'McGurk effect' where auditory perception is modified by visual information resulting in a misinterpretation of both modalities. As such, it can be said that incongruent sensory information may result in an interference effect. In a cross-modal illusion, results would be less accurate when compared to results from unimodal alone (McGurk & Macdonald, 1976).

Congruency of modalities in a CRT is important in ensuring an accurate interpretation can be made. In most patients, the visual modality would be congruent to the auditory modality but it is possible that in some patients, diminished or uncontrollable body movement may cause an incongruence of the visual modality compared to the auditory modality resulting in possible misinterpretation of cough responses. For example, patients with Parkinson disease have been found to have certain motoric features such as tremor at rest, rigidity and bradykinesia (Fahn, 2003). It is uncertain how these movements or lack of movements interfere with clinicians' judgement of cough responses. Additionally, patients' appearances might influence clinicians' judgements, with a patient who looks weak being judge to have a weak cough as opposed to a patient who is healthy looking. The possibility of cross-modal illusion may be tempered by familiarity with the patients' medical diagnosis and symptoms to avoid making a judgement error.

## **2.8 Aim of Study**

Past research has demonstrated that CRT can be used to assess the integrity of the cough reflex and help identify patients who are more likely to aspirate silently. However, to ensure that it is of clinical value, the reliability of CRT has to be determined. It is also of interest to know whether training in CRT would result in a higher reliability. As cross modality factors can influence perception of information, it would also be of interest to ascertain whether audiovisual modality has any advantage over the auditory or visual modalities in assisting in cough response interpretation. Information from the auditory and visual perception is gathered to understand types of information that cued clinicians in their decision making. As consensus methods can be used as a measure of relative validity (Fink, Kosecoff, Chassin, & Brook, 1984), ratings from individual clinicians were compared to consensus ratings by a panel of expert clinicians to assess the level of association. Thus, this study attempts to answer the following questions:

Question 1: How reliable are clinicians in rating cough response in CRT?

Question 2: Does training in CRT impact reliability?

Question 3: Which modality provides the most salient information?

Question 4: What information can be gleaned from the auditory and visual perception that may cue clinicians towards judging a cough response?

Question 5: What is the level of association between individual ratings by SLTs and consensus ratings by the expert team?

## **Chapter 3: Methodology**

### **3.1 Participants**

Inclusion criteria for the study required that participants were practicing SLTs who were actively involved with the diagnosis and management of dysphagia. All included participants reported normal or corrected-to-normal vision and have no hearing difficulties. As the study was completed through an online survey, all participants had access to a computer and internet. This study was approved by the University of Canterbury Human Ethics Committee low risks process.

### **3.2 Materials**

The audiovisual clips used in this study were taken from 10 hospitalised patients being administered a CRT in a clinical setting as part of their dysphagia swallowing assessment. All 10 patients gave their informed consent for recording and for use in research. Videos were recorded using a Mono HD flip video camera (CISCO, Irvine, CA) and edited first using iMovie (Apple, Cupertino, CA) then VideoPad Video Editor for extraction of visual-only, auditory-only and audio-visual clips. Consequently, the 10 patients' clips were made into 10 visual-only, 10 auditory-only and 10 audiovisual clips giving a total of 30 clips for rating.

Each audiovisual clip showed a patient receiving one 15-second dose of nebulised citric acid solution (diluted in 0.9% sodium chloride) through a facemask using a PulmoMate Compressor/ Nebuliser (Model 46501. DeVilbiss Healthcare LLC, Pennsylvania, US). In the visual-only clips, the audio signal was muted while in the auditory-only clips, only a black background was shown along with the audio signal. In all clips with audio, the mean

energy intensity of noise from the nebulizer was at a level between 50.9dB and 55.1dB with mean of 52.7dB. The 10 patients' clips were believed to be representative of the range of cough responses seen in the hospital setting.

### **3.3 Procedure**

All participants completed an online, computer based assessment consisting of three separate sessions. The first two sessions consisted of rating a counterbalanced mix of auditory-only and visual-only clips, while in the third session participants evaluated the audiovisual clips and answered a questionnaire. A more detailed description of each task is stated below.

#### **3.3.1 Online survey**

Participants were first informed that they were going to watch or listen to some auditory and/or video clips and were instructed to rate the cough response of each clip (see Figure 3). Before rating, participants were presented with a definition of cough (see Figure 4) and benchmarking stimulus, that is, clips of a cough versus no cough and a strong versus a weak cough. Definitions provided were taken from Morice et al. (2006, p. i5) and from previous studies by this lab.

In the first session, participants rated ten auditory-only or visual-only clips. The order of presentation for the 10 clips was set in which auditory-only alternates with visual-only clips or vice-versa. The 10 clips shown were a representation of auditory and visual segments from the ten different patients such that participants could not make an association between auditory-only and visual-only clip. Participants rated the presented stimulus for presence or absence of cough first, followed by strength of the cough (strong or a weak cough) only when



presence of a cough response was selected. Participants were also encouraged to provide a written justification for their ratings. After finishing the session, participants were only able to do the next series after a minimum delay of 48 hours.

Participants went through a similar process for the second and third sessions. In the second session, participants rated the next 10 auditory-only and visual-only clips and in the third series, participants rated 10 audiovisual clips. Altogether, 30 clips were rated.

### **3.3.2 Questionnaire**

All participants filled in an online questionnaire to provide general demographic data such as age, case load, experience level and country in which they worked. All participants were informed that the information from the questionnaire would be kept confidential. A sample of the questionnaire is provided in appendix 1.

### **3.4 Consensus Ratings by Expert Team**

All 30 clips were viewed by a panel of three clinicians who were considered experts in CRT to determine consensus ratings on the interpretation of cough. These clinicians are actively engaged in research on CRT and have been performing CRT on patients. The consensus ratings were compared with participants' ratings as a measure of relative validity.

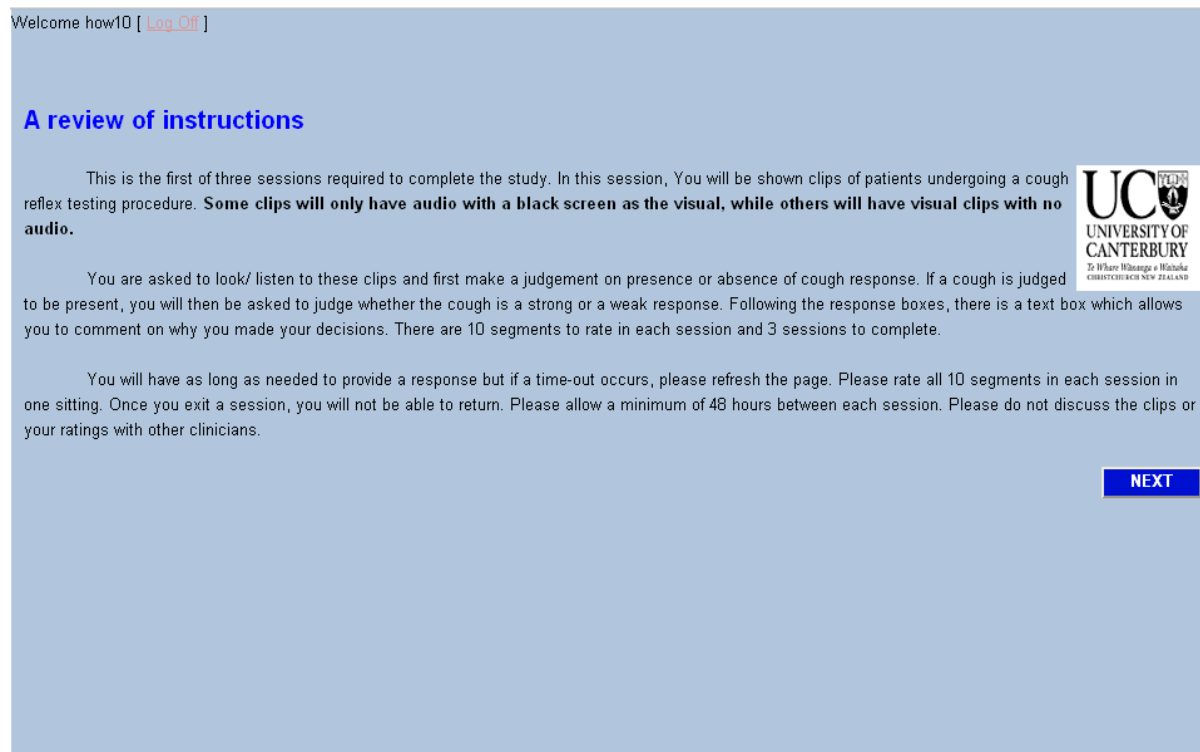


Figure 3. Instructions provided to participants in each session before ratings begin.

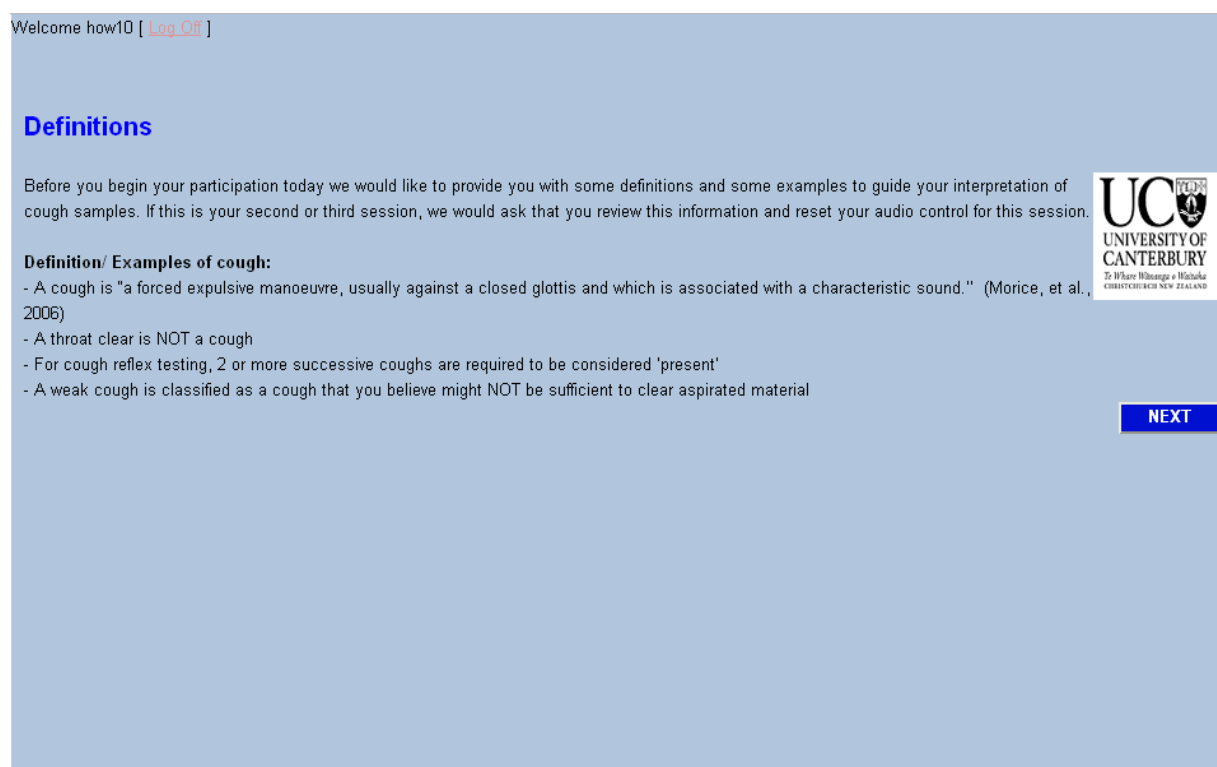


Figure 4. Definitions provided to participants in each session before ratings begin.

### **3.5 Data Analysis**

Data from the questionnaire and ratings were extracted from the online survey as a Microsoft excel spread sheet for analysis. Inter-rater agreement on ratings of cough response in auditory-only, visual-only and audiovisual clips were measured through Fleiss' Generalized Kappa using an online adaptable Microsoft excel spreadsheet (King, 2004). Participants were further divided into two groups depending on whether they had training in CRT. Inter-rater agreement was tabulated and compared. Level of association between ratings from participants and consensus ratings was performed using Pearson's chi-square test in SPSS 19.0. Comments from participants were compiled and analysed qualitatively by identifying common themes.

## Chapter 4: Results

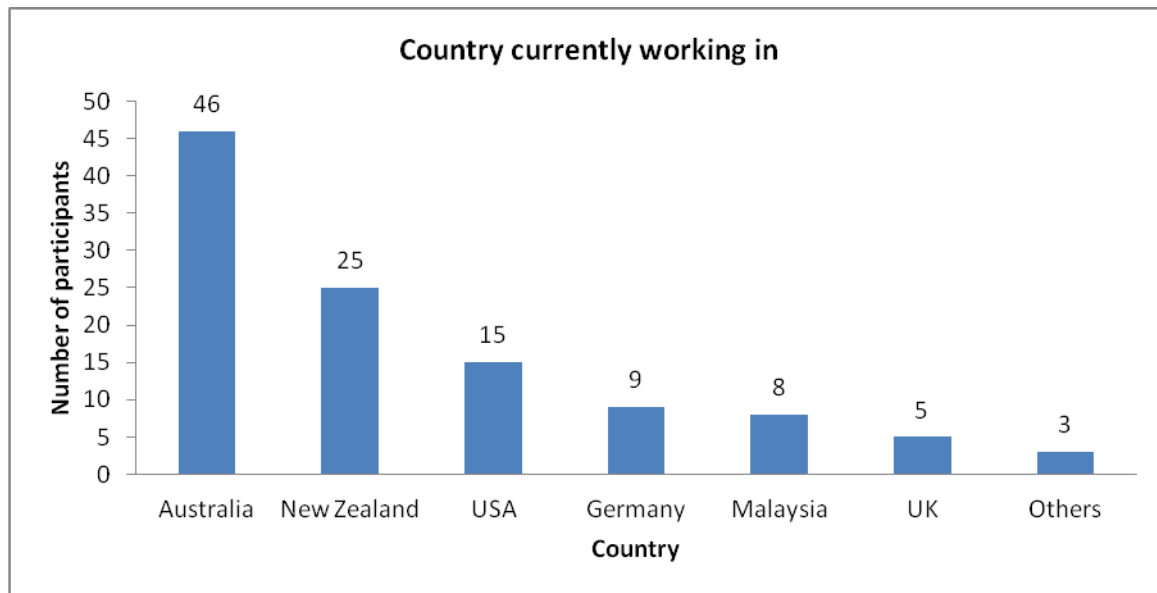
The results from this study are presented in four sections. Demographic data from the questionnaire is summarized and presented in the first section. In the second section, inter-rater agreement is reported. A comparison between individual ratings and consensus ratings is reported in the third section and the last section describes some of the common cues used by clinicians to assist decision making.

### 4.1 Demographic Data

A total number of 191 participants from nine different countries consented to participate in this study and 114 (59.7%) completed the study. Out of 114 participants, three were excluded as they did not meet the criteria of having normal or corrected-to-normal vision and hearing, leaving 111 participants. Participants who consented to participate but did not follow through with any tasks (25.7%), did not complete all of three series (11.0%), or pulled out (3.7%) were also excluded from the study. The country in which participants were currently working in is presented in Figure 5.

#### *Age of participants*

Out of 111 participants, the majority were aged between 26 – 35 years-old age group (55.9%). The rest of the participants were between 36 – 50 years-old group (21.6%), 51 – 75 years-old group (14.4%) and ‘less than 25 years-old group’ (7.2%). One participant did not indicate her/his age.



*Figure 5. Country in which participants were currently working.*

### *Work Settings*

All participants reported their work settings with many indicating that they worked in one setting only (68.4%). Other participants (31.6%) worked in two or more settings. When divided into the type of work settings, the majority of the participants reported working in acute hospitals (42.5%). Other participants worked in rehabilitation hospitals (19.3%), outpatient clinic (14.9%), home health care (7.7%), private practice (4.4%), chronic care (3.3%) and school (1.1%). Twelve participants (6.6%) stated that they worked in a setting other than those listed.

### *Years of experience as SLTs and in managing the dysphagic population*

The participants involved in this study were quite evenly represented in their levels of experience. Slightly more than a quarter (27.0%) of the participants reported at least 13 years

of experience working as a SLT. A roughly equal number of participants reported working for at least 1 to 3 years (20.7%) and 4 to 6 years (21.6%). Other participants reported 7 to 9 years of experience (13.5%) and 10 to 12 years of experience (11.7%). Only 5.4% of the participants reported working as a SLT for less than a year.

As with the level of experience as a SLT, the participants were also quite evenly represented in their levels of experience in managing patients with dysphagia. A quarter (25.2%) of the participants had at least 13 years of experience managing adult dysphagia with slightly less than a quarter with at least 1 to 3 years (23.4%) and 4 to 6 years (23.4%). Other participants reported 7 to 9 years of experience (14.4%), 10 to 12 years of experience (0.9%) and less than a year of experience (7.2%). Six participants (5.4%) did not report their years of working experience with the adult dysphagic population.

### *Dysphagia Case Load*

A total of 108 participants provided estimated percentages of adult dysphagic patients seen in their clinical practice (dysphagia caseload). The three participants who did not report had indicated that they have experience working with this population. Forty-three (38.7%) of the participants had a dysphagia case load of 75 to 100%, 28.8% had a dysphagia case load of 51 to 75% and 29.7% had a caseload of less than 50%.

### *Experience and training on CRT*

Only 21.6% of the participants reported that they were currently using CRT on patients at the time of the survey and of these, 54.2% had some form of training in CRT. The

majority were not using CRT as part of their dysphagia assessment (78.4%) although some of them had some form of training in CRT (9.9%). Table 2 represents the number of participants who did training in CRT and those who carried out this test.

Among participants who performed CRT, most of them had been carrying out CRT for less than a year (45.8%) while 25% of them had been using this test for 1 to 3 years. Only a quarter (25%) of the participants had experience of more than 4 years in performing CRT. Years of experience in performing CRT for one participant is unknown. Regardless, all 111 participants reported that they listen for cough sounds while assessing patients with dysphagia indicating they have some experience or awareness of cough assessment.

Table 2

*Number of participants that are performing CRT and trained in CRT*

<b>Performing CRT</b>	<b>Training in CRT</b>		<b>Total</b>
	<b>Yes</b>	<b>No</b>	
<b>Yes</b>	11	13	24
<b>No</b>	11	76	87
<b>Total</b>	22	89	111

## 4.2 Inter-rater Agreement

Fleiss' Generalized Kappa was used to analyse inter-rater agreement between the 111 participants for the ratings of cough in the auditory-only, visual-only and audiovisual clips. The interpretation of strength of agreement used for  $\kappa$  values were based on suggestions by Landis and Koch (1977) and is presented in Table 3. Overall inter-rater agreement for the judgement of presence versus absence of cough response in the auditory-only and audiovisual modality was .75, 95% CI [0.74, 0.75] and .76, 95% CI [0.76, 0.77] respectively indicating substantial agreement. Inter-rater agreement in the visual-only modality was slightly lower ( $\kappa$  = .65, 95% CI [0.65, 0.66]) though still indicating substantial agreement.

Inter-rater agreement was also tabulated for participants who graded the cough to either 'weak' or 'strong' when a cough response was judged to be present. Only responses from eight clips were analysed in the audiovisual and auditory-only modality as in the remaining two clips, all participants rated the cough responses as absent. Inter-rater agreements on the grading were found to be fair on both auditory-only ( $\kappa$  = .24, 95% CI [0.24, 0.25]) and audiovisual modality ( $\kappa$  = .25, 95% CI [0.24, 0.25]) whereas inter-rater agreement was slight in the visual-only modality ( $\kappa$  = .18, 95% CI [0.17, 0.19]). Kappa values with the corresponding confidence intervals and percent of agreement for all modalities and grades are presented in Table 4.

The data from the audiovisual modality was further divided into clinicians who had training in CRT versus clinicians with no training. Fleiss' Generalized Kappa was performed for inter-rater agreement in these groups. In both groups, the inter-rater agreements for the presence of a cough response were substantial. Those with training had inter-rater agreement of .74 (95% CI [0.70, 0.78]) and untrained .77 (95% CI [0.76, 0.78]). However, grading of cough strength in both groups were only fair with those trained having a score of .21 (95% CI



[0.17, 0.25] and untrained having a score of .25 (95% CI [0.24, 0.26]). It should be noted that only responses for eight clips were analysed as in the remaining two clips, all participants rated the cough responses as absent. **Table 5** reports the inter-rater agreement for these two groups.

Table 3

*Interpretation of strength of agreement based on Kappa values. Reprinted from “The Measurement of Observer Agreement for Categorical Data,” by J.R. Landis and G.G. Koch, 1977, Biometrics, 33, p.165. Reprinted with permission.*

Kappa Statistics	Strength of Agreement
<0.00	Poor
0.00-0.20	Slight
0.21-0.40	Fair
0.41-0.60	Moderate
0.61-0.80	Substantial
0.81-1.00	Almost Perfect

Table 4

*Overall inter-rater agreement of cough response in auditory-only, visual-only and audiovisual modality*

Grade	Modality	Generalized Kappa ( $\kappa$ )	Confidence Interval	Percent Agreement (pa)
Absent/ Present	Auditory-only	.75	0.74 to 0.75	.89
	Visual-only	.65	0.65 to 0.66	.86
	Audiovisual	.76	0.76 to 0.77	.90
Strong/ Weak	Auditory-only	.24	0.24 to 0.25	.53
	Visual-only	.18	0.17 to 0.19	.39
	Audiovisual	.25	0.24 to 0.25	.53

Table 5

*Inter-rater agreement of cough response as observed through level of training*

Groups	Generalized	Confidence	Percent
	Kappa ( $\kappa$ )	Interval	Agreement (pa)
Ratings of Absent or Present			
Trained (n= 22)	.74	0.70 to 0.78	.89
No training (n=89)	.77	0.76 to 0.78	.90
Ratings of Weak or Strong			
Trained (n= 22)	.21	0.17 to 0.25	.50
No training (n=89)	.25	0.24 to 0.26	.54

### 4.3 Comparison of Participants' Ratings with Consensus Ratings

Pearson's chi-square test was used to determine the relationship between ratings done by participants with consensus ratings by the expert team. An additional Cramer's V statistic was used to determine the strength of association. A significant association was observed in the auditory-only modality on the assigned ratings of absent versus present between the two groups of raters  $\chi^2 (1) = 477.92, p < .001$ . The results also showed that there is a large association between participants' ratings with consensus ratings on presence or absence of cough responses in the auditory-only modality ( $\phi_c = .66$ ).

Excluding one missing rating, a total of 1109 ratings from the participants on absent versus present cough responses were compared to consensus ratings made by the experts in the visual-only modality. A significant association was also observed in the visual-only modality on the assigned ratings between the two groups of raters  $\chi^2 (1) = 433.99, p < .001$ . A large association is observed between participants' ratings with consensus ratings on rating cough responses in the visual-only modality ( $\phi_c = .63$ ).

In the audiovisual modality, all participants rated with a total of 1110 ratings. A significant association was found between participants' and consensus ratings on the assigned ratings of absent versus present response  $\chi^2 (1) = 688.31, p < .001$  with a large association observed between participants' ratings with consensus ratings ( $\phi_c = .79$ ). **Table 6** displays the ratings made by participants on grading of absence and presence of cough in comparison with consensus ratings.

Responses of weak and strong from participants who rated cough response as present were compared with consensus ratings. A total of 694 ratings on the auditory-only modality were analysed using Pearson's chi-square test with a significant association found between

the two groups of raters on the assigned ratings  $\chi^2 (1) = 206.98$ ,  $p < .001$ . A large association was observed between the two groups of raters in rating for strength of cough ( $\phi_c = .55$ ).

A total of 630 ratings were compared to consensus ratings for ratings on strength of cough in the visual-only modality. A significant association was also seen in the visual-only modality on the assigned ratings between the two groups of raters  $\chi^2 (1) = 122.47$ ,  $p < .001$ . However, the level of association is lower as compared to the auditory-only modality ( $\phi_c = .44$ ).

In the audiovisual modality, 662 ratings were analysed. A significant association was observed on the assigned ratings between the two groups of raters  $\chi^2 (1) = 32.45$ ,  $p < .05$ . There was only small to medium association between the two groups of raters on the assigned ratings ( $\phi_c = .22$ ). Table 7 displays the ratings made by participants on strength of cough in comparison with consensus ratings.

Table 6

*Comparison of ratings of presence of cough by participants with consensus ratings*

		Participants' ratings	
		Absent	Present
Auditory-only (n=1110)			
Consensus ratings	Absent	255	78
	Present	83	694
Visual-only (n=1109)			
	Absent	275	169
	Present	32	633
Audiovisual (n= 1110)			
	Absent	330	114
	Present	4	662

Table 7

*Comparison of ratings of strength of cough by participants with consensus ratings*

		Participants' ratings	
		Weak	Strong
Auditory-only (n=694)			
Consensus ratings	Weak	319	155
	Strong	19	201
Visual-only (n=630)			
	Weak	281	135
	Strong	45	169
Audiovisual (n=662)			
	Weak	241	200
	Strong	69	152

#### 4.4 Cues That Guide Clinicians toward Decision Making

Comments to the question “Please comment on why you made this decision” were collected and analysed. Repeated key-words and phrases were identified as the most common cues applied by clinicians in decision making. These cues include those provided initially in the definitions that guide clinicians in their interpretations.

##### *Perceived ability to clear secretions or aspirated materials*

In the presence of a cough, participants would determine the strength of cough by their perceived ability of the cough to clear any aspirated material. Words such as *productive* and *effective* were also used to describe a strong cough. One of the participants commented:

*‘Although I could hear the coughs, unlike some of the other samples, I had to question whether it rated strong enough to expel aspirated material. Therefore I rated it as weak’.*

##### *Intensity of cough sounds*

Participants used a wide variety of words to indicate the intensity of cough sounds. Some of these words were *loud*, *soft*, *low volume*, *explosive* and *forceful*. Strong coughs were described to be *loud in volume* and *explosive* while weak coughs were the opposite. An example of a participant commenting on the intensity of cough sound:

*“Sounded very weak, not forceful and was not loud”*



### *Quantification of cough number*

One of the criteria used in CRT is that at least two or more successive cough must be heard or observed to be considered as ‘present’. Many participants commented on the number of cough observed though the numbers may not be similar among participants listening or watching the same clip. Examples of two comments by different participants on the same clip:

*‘I only heard one cough. This does not meet your criteria of at least 2 successive coughs.’*

(Participant 1)

*‘There is 2x coughs which meets the research criteria of a cough’.* (Participant 2)

### *Quality of cough sounds and other vocalization*

The quality of cough sounds and other vocalization sounds were used in an attempt to determine the grade of cough. Words used frequently by participants to include *wet*, *wheeze/wheezy* and *breathy*. One participant commented:

*‘This cough sounded weak due to its wheezy nature - I do not feel confident this patient could clear aspirated material if required to do so.’*

Many participants also commented on whether the response from CRT was a cough or other sounds. Participants described some response as *not a cough* but *throat clear*, *sneezing*, *gagging* and *vocalizations*. One such example by a participant:

*‘... I would rate this as a throat clear rather than a cough but I did contemplate rating as a weak cough however I was led by the fact that it was a minimal response’*

### *Body movements associated with coughing*

Descriptions of movements were used by participants while rating visual-only clips. Chest and trunk (body) movement were often used to decide the grade of cough response. A strong cough was judged to be present when accompanied by forceful and strong body movements and weak or absent if minimal or no body movement was observed. Presence of cough was also determined by any changes in respiration. Other movements which cued participants include degree of mouth opening and level of movements of the head, neck and shoulders. However, many participants stated that in order to determine whether a cough response is strong or weak, auditory information would be needed. As noted by one participant:

*'I believe there is a cough despite nil audio. There is purposeful inspiration hold and then air release on multiple occasions. I would hypothesis that the cough may be reduced in volume due to the lack of deep inspiration, minimal chest drop post cough, nil jaw drop/mouth opening to quickly release the air. I would need to listen though to check my hypothesis of weak cough. I do not feel that I can accurately determine based on vision alone.'*

### *Patients' reaction and appearance*

Participants' decisions were influenced by the appearances of patients and their reaction towards the tussive agents. Participants commented on observable facial grimacing, signs of distress and general appearance of weakness. Signs of discomfort were more likely to be associated with a strong cough while appearances of frailness were associated with weak cough. This would probably bias the clinician towards decision making. Two such comments by participants while watching a visual-only clip and an audiovisual clip:

*'Again a borderline decision: multiple coughs appearance of adequate laryngeal and thoracic movement. Unable to judge VF adduction without sound but the pt appeared frail and distressed which would negatively influence my decision to offer PO trials.'* (Participant 1: visual-only clip)

*'level of visible distress impacting on perception of strength of cough despite audibility'*  
(Participant 2: audiovisual clips)

## **Chapter 5: Discussion**

This is the first study to evaluate inter-rater reliability of subjective cough judgements in a large population of clinicians. The study also specifically evaluated the independent contributions of auditory and visual modalities to those judgments. A comparison with consensus ratings provided a relative measure of clinicians' accuracy in judging cough responses.

### **How reliable are clinicians in rating cough response in CRT?**

Clinicians performing CRT in a real life situation would use both their visual and auditory senses to observe for cough responses. As such, inter-rater agreement in the audiovisual modality represents a normal situation. Inter-rater agreement in deciding presence or absence of cough was substantial ( $\kappa=.76$ ) while it was only fair ( $\kappa=.25$ ) in deciding strength of cough. This indicates that clinicians were able to detect the presence of a cough relatively well but agreement is much reduced when asked to judge the strength of a cough in CRT.

The results from this study were dissimilar to Chang et al.'s (2005) study whereby both wet/dry cough and brassy/non-brassy cough had similar levels of reliability with both rated as very good. One possible reason for the differences may be because clinicians were measuring different aspects of cough. In Chang et al.'s (2005) study, only quality of cough was measured but in the present study, the measurement on grading of cough may include both quality and intensity of cough. When SLTs were asked how they decide on grading, words such as wet, breathy, low volume were used. The added factor may add to clinicians' confusion thus result in lower inter-rater agreement.

Differences were also observed when comparing grading of presence versus absence of cough and strength of cough. Judging for presence versus absence of cough response is a dichotomous decision with no underlying continuum whereas strength of response, although presented as a binomial data, is actually based on an underlying continuum. As mentioned by Irwig and Groeneveld (1988, p. 956) in radiological assessments “readers may use different thresholds due to differences in their visual perception or decision attitude, even in the presence of criteria which attempt to define clear boundaries between categories”. This statement would apply in the current situation as well with the presence of response bias when judging strength of cough due to clinicians having their own threshold choice. Indeed similar results were found in reliability studies on VFSS. Reliability on presence versus absence of aspiration was higher when compared to whether laryngeal elevation is normal, as this is based on an underlying continuum with restricted definition (Kuhlemeier, et al., 1998).

In summary, the reliability of clinicians in judging cough response in CRT varies according to the trait being rated. Clinicians are better in judging for the presence of cough but are not very reliable when deciding the strength of cough. It is uncertain whether grouping the grades differently as in Addington group’s (Addington, et al., 2005; Addington, Stephens, Gilliland, et al., 1999; Addington, Stephens, & Gilliland, 1999), studies would result in better reliability. In the current study, the grades of cough were grouped into present versus absent and strong versus weak, while in Addington group’s (Addington, et al., 2005; Addington, Stephens, Gilliland, et al., 1999; Addington, Stephens, & Gilliland, 1999) studies, grades were grouped into normal versus abnormal (weak and absent) with significantly consistent results. Further research is needed looking into whether reliability is improved with different groupings.

### **Does training in CRT impact reliability?**

Clinicians who had some form of training in CRT performed no better than those who were untrained. Although no studies have specifically looked at cough response in CRT, the results obtained here are in agreement with other reliability studies (Borr, et al., 2007; Leslie, et al., 2004) whereby training and experience does not necessarily produce higher inter-rater agreement. However, this does not imply that training has no effect at all in improving reliability. Among clinicians who reported that they have some form of training in CRT, type and level of training were not specified. The amount of training might be an important variable rather than just presence or absence of training.

It is possible that training to criteria may improve the reliability in ratings. This has been found in a study on VFSS ratings, whereby higher inter-rater agreement was observed when clinicians were able to discuss rating scales and issues arising from it in groups prior to individual ratings (Scott, et al., 1998). Poorer agreement was seen when clinicians rated independently relying only on a scale as a guide without any prior discussions (Scott, et al., 1998).

In summary, although training was not seen to have an effect on reliability in this study, it is possible that more specific training needs to be conducted before an effect on reliability is seen. A discussion among clinicians on cough response criteria prior to the beginning of the study is likely to increase the inter-rater agreement among clinicians.

### **Which modality provides the most salient information?**

Although clinicians relied on both auditory and visual modality in rating cough responses, there has been no research to show that audiovisual modality provides the most salient information in CRT. Past research in speech comprehension however suggested that

cross-modal enhancement would occur in a congruent multimodal situation resulting in reduction of ambiguity (Arnold & Hill, 2001; Arnold & Oles, 2001). It was hypothesized that inter-rater agreement for cough responses would be higher in audiovisual modality as compared to unimodal (auditory and visual) due to reduction in ambiguity.

Surprisingly, the results in this study disagree with past research of cross-modal enhancement in non-speech sounds. Inter-rater agreements for ratings of cough response were similar when judged in the audiovisual modality and auditory modality. The audiovisual modality did not provide any extra information as compared to auditory alone. The ceiling effect found in Schmid and Ziegler's (2006) study does not appear to be a reason for the lack of cross-modal enhancement. In their study, the error rates were already low preventing further enhancement. Granted that normal subjects appeared to have higher error rates when doing bimodal non-speech discrimination tasks as compared to bimodal speech discrimination tasks but cross-modal enhancement should still occur (Eramudugolla, et al., 2011). In this current CRT study, although the percent of agreement in ratings for presence of cough is high (90%), the ratings for strength are only about 50% allowing possible enhancement which was not observed.

Another more probable explanation for the similar results is that ratings of cough response are more heavily influenced by decisional related process rather than perceptual processing. Previous studies on perception of non-speech tasks mostly required participants to detect the presence of a sound (Eramudugolla, et al., 2011) or discriminate sounds (Schmid & Ziegler, 2006). For example, in Schmid and Ziegler's (2006) study, participants were presented with pairs of visual-only, auditory-only, audiovisual, or auditory-only followed by visual-only clips, and asked to judge whether the clips were the same. As such, it relies more heavily on perception and does not require much cognitive skills. In contrast to the situation in CRT, clinicians would only judged cough to be present when at least two consecutive

cough sounds were heard or observed. Clinicians also had to determine whether the responses heard or observed were a cough as opposed to throat clearing, sneezing or vocalizations. If a clinician judged a cough to be present, further decision has to be made on the strength of cough, which as has been mentioned earlier, not a true dichotomous decision. Adding to the awareness that there is a high cost for misclassification, clinicians would be more cautious in their decision making. As such, clinicians relied heavily on cognitive skills for their decision making. It is uncertain how the decisional-related process interacts and influence perceptual processing and whether this would reduce any cross-modal enhancements.

As expected, the inter-rater agreement in the visual modality was the lowest among the three modalities. Presence of a response may be possible to determine through visual perception alone but it would be difficult to identify whether the response is a cough response or other vocalizations and the strength of cough. Furthermore, the visual information was degraded by the presence of a face mask that covered the mouth and nose, and in certain patients, a blanket that covered the patient's body. This would probably account for the lower agreement in the visual modality.

In summary, no advantage was found in the audiovisual modality as compared to auditory-only modality. Although it might be easier for clinicians to make judgements with the extra visual information from audiovisual modality, clinicians should be aware of the possibility of bias which might result in reduce accuracy. This is discussed in further detail in the section of relative validity. The visual-only modality in itself is insufficient for clinicians to make a judgement. As this study was not designed to separate perceptual processing from decisional related processing, further studies are necessary to investigate how one affects the other.



### **What information can be gleaned from the auditory and visual perception that may cue clinicians towards judging a cough response?**

An attempt was made to understand how visual and auditory information cues clinicians in their decisions. Understandably, most clinicians made use of the definitions provided. Clinicians differentiated it with other vocalizations, commented on the intensity and perceived ability of the cough to clear secretions or aspirated materials, described the quality of cough sound and quantify the number of coughs heard. However, it was noted that there was individual variation in how clinicians perceive coughs due to individuals' own threshold choice.

It was expected that clinicians would describe the quality of cough sound heard as explosive or productive as it was stated in the definition. Besides these terms, a few other terms were used such as breathy, wet and wheezy. Wet cough has been associated with cough with mucus (Smith, Ashurst, et al., 2006), which might indicate inability to expectorate mucus and hence clinicians may rate it as weak. It is unsure how clinicians relate wheezy cough with grading of strength of cough. It is possible that as wheezing has been associated with foreign body aspiration (Paintal & Kushner, 2007), clinicians were more wary of making a judgement of strong. Dysphonia which highly indicates vocal fold pathology may cue clinicians that full glottis closure is not possible and therefore reduce cough efficacy.

The presence of body movement associated with cough was mentioned but not elaborated by Smith, Earis, et al. (2006). Based on description of musculature activity during cough (Fontana, 2008; Fontana & Lavorini, 2006), it was assumed that the chest movement and mouth opening would be seen during cough and this was observed in comments made by clinicians. Some other movements which were described by clinicians include head, neck, shoulders and body movement. During coughing various muscles such as diaphragm,

abdominal muscles, intercostals muscles and the glottis are activated and level of muscles activation increased as the cough intensity increased (Chang, 1999; Fontana, 2008; Fontana & Lavorini, 2006). This may translate into more vigorous movement observed in a high intensity cough. Not only the chest would be moving but also the whole body including the head and neck. Comments from clinicians appeared to support this notion, whereby cough was more frequently described as strong when it is accompanied by rigorous body movement.

One interesting detail that emerged from this study is the response bias by clinicians. Some clinicians were influenced by the appearance of the patient and their response to citric acid. A patient who looked weak and frail was more likely to be judged having a weak cough by clinicians. Clinicians also appeared to equate levels of distress with strength of cough. A patient who looked highly distressed was more likely to be judged having a strong cough. This might not be true, as a patient could appear very distressed but having a weak cough indicating possible intact sensory but impaired motor components of reflexive cough. This would result in patient being irritated by the citric acid but was ineffective in producing a strong cough to clear it. Thus far, response bias has not been described in CRT studies though studies on the impact of physical appearance on clinical judgement have been described in other areas such as obesity. In a review article by Puhl and Brownell (2001), the investigators described how obese individuals were viewed negatively by health care professionals influencing their clinical judgement, diagnosis and care. Physicians were reluctant to treat obese individuals as they do not expect them to succeed, viewing these individuals as weak-willed (Maddox and Liederman in Puhl & Brownell, 2001). However, unlike physicians' attitude with obesity, SLTs may be justifiable in being careful with frail patients. The elderly population with lower functional status and multiple medical conditions are known to be at a higher risk for acquiring aspiration pneumonia especially when coupled with oropharyngeal

dysphagia (Cabre et al., 2010; Langmore, et al., 1998). Further research in this area is needed to ascertain how bias affects ratings and the consequence of bias in the management of patients with dysphagia.

In summary, clinicians gather various information through auditory and visual perception modalities and use this information to assist them in rating cough responses. Oftentimes, clinicians have their own strategy on signs to observe and would consistently use them in their ratings. It might be beneficial to have a discussion among clinicians on which information is the most valid in assisting ratings of cough.

#### **What is the level of association between individual ratings by SLTs and consensus ratings by the expert team?**

Consensus methods have been used as a measure of relative validity (Fink, et al., 1984). As such, in the absence of a more objective measurement to validate the results, comparison with expert consensus would give an indication on accuracy of ratings. Overall, a significant association was observed on ratings of presence of cough and strength of cough in all modalities (visual-only, auditory-only and audiovisual) between the two groups of raters. However, the levels of association differ between modalities and between traits.

Clinicians were able to rate with better accuracy through the audiovisual modality for presence of cough but not when rating strength of cough. Strength of cough was rated with better accuracy when using auditory-only perception. A possible reason for it might be that clinicians were biased by the visual information obtained through the audiovisual modality while rating for strength of cough thus resulting in lower accuracy. As stated earlier, clinicians were swayed by the appearances of patients and patients' response to CRT, with

frail looking patients being assigned a weak cough. This bias would not be present when rating presence or absence of cough as the criteria is clearer.

In summary, individual ratings agreed with the consensus ratings, providing a relative measure of validity. Clinicians were more accurate when rating for presence of cough when compared to strength of cough. Possible biasness occurs when clinicians rate strength of cough through audiovisual modality. As it is unrealistic to ask clinicians not to look at patients while performing CRT, clinicians should be made aware of the possibility of bias and reduced accuracy, and take precautions to avoid it.

## **Limitations**

There were a few limitations to this study. The study was designed as an online study with the purpose to recruit as many clinicians as possible as participants. However, there were many uncontrollable factors in online studies design, including compliance of participants. Although instructions were provided at the beginning of each session, it is uncertain whether participants were compliant in following it, such as maintaining the level of loudness throughout all clips. There was also difficulty in controlling time length for participants to complete all three series. A minimum of 4 days was needed to complete the whole study and some participants adhered strictly to it while the majority took more than 4 days to complete the survey. As the purpose of the break was to allow the participants to ‘forget’ their earlier decision, avoiding any memory bias, it is considered acceptable to have participants to have a longer gap in between series. However, it is uncertain whether time factor would affect the results.

Another difficulty was in randomizing the clips for participants. Preferably, each participant would rate clips in different orders but unfortunately this could not be done. As such, the solution was to counterbalance the order of clips. About half of the participants rated the first-series of 10 unimodal clips first, while the other half of the participants started with ratings of the second-series of 10 unimodal clips. This would hopefully reduce any exposure bias. However, it should be noted that both groups did the audiovisual clips in the last series. The reason for it is that audiovisual clips may be more memorable to participants as compared to unimodal clips and therefore would result in memory bias. As such, it was of the opinion that avoiding memory bias is more important and this outweighs the training effect.

## **Conclusion**

This study provided valuable information on how reliable and accurate clinicians are in interpreting cough responses from CRT. It also looked into factors that contribute to reliability of interpreting cough responses. Clinicians are more reliable and accurate at rating presence or absence of cough than strength of cough. A mixture of auditory and visual cues is used to assist in decision making though clinicians need to be aware of the possibility of response bias, especially when utilizing the visual information. Training was not shown to have an effect on reliability. It would be informative to know whether reliability could be improved by having discussion and training to criterion prior to ratings. As the focus of this study was on reliability and its influencing factors, only relative measure of validity was obtained. Future research is needed to look into validation by comparing subjective interpretation of cough responses with objective measurement such as airflow rate and volume.

## References

- Addington, W. R., Stephens, R., Widdicombe, J., & Rekab, K. (2005). Effect of stroke location on the laryngeal cough reflex and pneumonia risk. *Cough*, 1(1), 4.
- Addington, W. R., Stephens, R. E., Gilliland, K., & Rodriguez, M. (1999). Assessing the laryngeal cough reflex and the risk of developing pneumonia after stroke. *Archives of Physical Medicine and Rehabilitation*, 80(2), 150-154. doi: 10.1016/s0003-9993(99)90112-0
- Addington, W. R., Stephens, R. E., & Gilliland, K. A. (1999). Assessing the laryngeal cough reflex and the risk of developing pneumonia after stroke: an interhospital comparison. *Stroke*, 30(6), 1203-1207.
- Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour. *Seeing and Perceiving*, 23(1), 3-38. doi: 10.1163/187847510x488603
- Arnold, P., & Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology*, 92, 339-355.
- Arnold, P., & Oles, J. (2001). Bisensory augmentation of complex spoken passages. *British Journal of Audiology*, 35, 53-58.
- Bernstein, L. E., Auer Jr, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication*, 44(1-4), 5-18. doi: 10.1016/j.specom.2004.10.011
- Bickerman, H. A., Barach, A. L., & Drimmer, F. (1954). The experimental production of cough in human subjects induced by citric acid aerosols. preliminary studies on the evaluation of antitussive agents. *The American Journal of the Medical Sciences*, 228(2), 156-163.

- Bolser, D. C., Davenport, P. W., Golder, F. J., Baekey, D. M., Morris, K. F., Lindsey, B. G., & Shannon, R. (2008). Neurogenesis of cough. In K. F. Chung, J. G. Widdicombe & H. A. Boushey (Eds.), *Cough: Causes, mechanisms and therapy* (pp. 173-180): Wiley-Blackwell.
- Borr, C., Hielscher-Fastabend, M., & Lücking, A. (2007). Reliability and validity of cervical auscultation. *Dysphagia*, 22(3), 225-234. doi: 10.1007/s00455-007-9078-3
- Cabre, M., Serra-Prat, M., Palomera, E., Almirall, J., Pallares, R., & Clavé, P. (2010). Prevalence and prognostic implications of dysphagia in elderly patients with pneumonia. *Age and Ageing*, 39(1), 39-45. doi: 10.1093/ageing/afp100
- Callan, D. E., Callan, A. M., Kroos, C., & Vatikiotis-Bateson, E. (2001). Multimodal contribution to speech perception revealed by independent component analysis: a single-sweep EEG case study. *Cognitive Brain Research*, 10(3), 349-353. doi: 10.1016/s0926-6410(00)00054-9
- Chang, A. B. (1999). Cough, cough receptors, and asthma in children. *Pediatric Pulmonology*, 28(1), 59-70.
- Chang, A. B. (2006). The physiology of cough. *Paediatric Respiratory Reviews*, 7(1), 2-8. doi: 10.1016/j.prrv.2005.11.009
- Chang, A. B., Gaffney, J., Eastburn, M., Faoagali, J., Cox, N., & Masters, I. (2005). Cough quality in children: a comparison of subjective vs. bronchoscopic findings. *Respiratory Research*, 6(1), 3.
- Chung, K. F., Bolser, D., Davenport, P., Fontana, G., Morice, A., & Widdicombe, J. (2009). Semantics and types of cough. *Pulmonary Pharmacology & Therapeutics*, 22, 139-142.



- Colodny, N. (2002). Interjudge and intrajudge reliabilities in fiberoptic endoscopic evaluation of swallowing (fees) using the penetration–aspirations scale: A replication study. *Dysphagia*, 17(4), 308-315. doi: 10.1007/s00455-002-0073-4
- Cox, I. D., Wallis, P. J., Apps, M. C., Hughes, D. T., Empey, D. W., Osman, R. C., & Burke, C. A. (1984). An electromyographic method of objectively assessing cough intensity and use of the method to assess effects of codeine on the dose-response curve to citric acid. *British Journal of Clinical Pharmacology*, 18(3), 377-382.
- Davis, C., & Kim, J. (2004). Audio-visual interactions with intact clearly audible speech. *The Quarterly Journal of Experimental Psychology* 57(6), 1103-1121. doi: 10.1080/02724980343000701
- Dicpinigaitis, P. V. (2003). Short- and long-term reproducibility of capsaicin cough challenge testing. *Pulmonary Pharmacology & Therapeutics*, 16(1), 61-65. doi: 10.1016/s1094-5539(02)00149-9
- Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, 1242(0), 136-150. doi: 10.1016/j.brainres.2008.03.071
- Eckert, D. J., Catcheside, P. G., Stadler, D. L., McDonald, R., & et al. (2006). Acute sustained hypoxia suppresses the cough reflex in healthy subjects. *American Journal of Respiratory and Critical Care Medicine*, 173(5), 506-511.
- England, R., Martin, D. F., Morris, J., Neill, P. A., Park, C., Renwick, D. S., . . . Wyatt, R. (1998). Can bedside assessment reliably exclude aspiration following acute stroke? [Article]. *Age and Ageing*, 27(2), 99-106.
- Eramudugolla, R., Henderson, R., & Mattingley, J. B. (2011). Effects of audio-visual integration on the detection of masked speech and non-speech sounds. *Brain and Cognition*, 75(1), 60-66.

- Fahn, S. (2003). Description of parkinson's disease as a clinical syndrome. *Annals of the New York Academy of Sciences*, 991(1), 1-14. doi: 10.1111/j.1749-6632.2003.tb07458.x
- Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus methods: characteristics and guidelines for use. *American Journal of Public Health*, 74(9), 979-983.
- Fontana, G. A. (2008). Motor mechanisms and the mechanics of cough. In K. F. Chung, J. G. Widdicombe & H. Boushey (Eds.), *Cough: Causes, mechanisms and therapy* (pp. 193-205): Wiley-Blackwell.
- Fontana, G. A., & Lavorini, F. (2006). Cough motor mechanisms. *Respiratory Physiology & Neurobiology*, 152(3), 266-281.
- Fontana, G. A., & Widdicombe, J. (2007). What is cough and what should be measured? *Pulmonary Pharmacology & Therapeutics*, 20(4), 307-312.
- Fuller, R. W. (2002). Cough provocation tests: their clinical value. *Pulmonary Pharmacology & Therapeutics*, 15(3), 273-276.
- Garon, B. R., Sierzant, T., & Ormiston, C. (2009). Silent aspiration: results of 2,000 video fluoroscopic evaluations. *Journal of Neuroscience Nursing*, 41(4), 178-185
- Helfer, K. S. (1997). Auditory and auditory-visual perception of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 40(2), 432-443.
- Hiorns, M., & Ryan, M. (2006). Current practice in paediatric videofluoroscopy. *Pediatric Radiology*, 36(9), 911-919. doi: 10.1007/s00247-006-0124-3
- Irwig, L. M., & Groeneveld, H. T. (1988). Exposure—response relationship for a dichotomized response when the continuous underlying variable is not measured. *Statistics in Medicine*, 7(9), 955-964. doi: 10.1002/sim.4780070907
- Kelsall, A., Decalmer, S., Webster, D., Brown, N., McGuinness, K., Woodcock, A., & Smith, J. (2008). How to quantify coughing: correlations with quality of life in chronic

- cough. *European Respiratory Journal*, 32(1), 175-179. doi: 10.1183/09031936.00101307
- King, J. E. (2004). *Software solutions for obtaining a kappa-type statistic for use with multiple raters*. Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX.
- Kuhlemeier, K. V., Yates, P., & Palmer, J. B. (1998). Intra- and interrater variation in the evaluation of videofluorographic swallowing studies. *Dysphagia*, 13(3), 142-147. doi: 10.1007/pl00009564
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Langmore, S. E., Terpenning, M. S., Schork, A., Y., C., Murray, J. T., Lopatin, D., & Loesche, W. J. (1998). Predictors of aspiration pneumonia: how important is dysphagia. *Dysphagia*, 13, 69-81.
- Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*, 249(5), 413-422. doi: 10.1046/j.1365-2796.2001.00812.x
- Lee, P. C. L., Cotterill-Jones, C., & Eccles, R. (2002). Voluntary control of cough. *Pulmonary Pharmacology & Therapeutics*, 15(3), 317-320.
- Leslie, P., Drinnan, M. J., Finn, P., Ford, G. A., & Wilson, J. A. (2004). Reliability and validity of cervical auscultation: A controlled comparison using video fluoroscopy. *Dysphagia*, 19(4), 231-240. doi: 10.1007/s00455-004-0007-4
- Marik, P. E. (2001). Aspiration pneumonitis and aspiration pneumonia. *The New England Journal of Medicine*, 344, 665-671.

- Masiero, S., Pierobon, R., Previato, C., & Gomiero, E. (2008). Pneumonia in stroke patients with oropharyngeal dysphagia: a six-month follow-up study. *Neurological Sciences*, 29(3), 139-145. doi: 10.1007/s10072-008-0925-2
- McCullough, G. H., Wertz, R. T., & Rosenbek, J. C. (2001). Sensitivity and specificity of clinical/bedside examination signs for detecting aspiration in adults subsequent to stroke. *Journal of Communication Disorders*, 34(1-2), 55-72.
- McCullough, G. H., Wertz, R. T., Rosenbek, J. C., Mills, R. H., Webb, W. G., & Ross, K. B. (2001). Inter- and intrajudge reliability for videofluoroscopic swallowing evaluation measures. *Dysphagia*, 16(2), 110-118. doi: 10.1007/s004550010004
- McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. [10.1038/264746a0]. *Nature*, 264(5588), 746-748.
- Midgren, B., Hansson, L., Karlsson, J.-A., Simonsson, B. G., & Persson, C. G. A. (1992). Capsaicin-induced cough in humans. *American Journal of Respiratory and Critical Care Medicine*, 146(2), 347-351. doi: 10.1164/ajrccm/146.2.347
- Morice, A. H. (2008). Rebuttal: Cough is an expiratory sound. *Lung*, 186(0), 7-9. doi: 10.1007/s00408-007-9039-5
- Morice, A. H., Fontana, G. A., Belvisi, M. G., Birring, S. S., Chung, K. F., Dicpinigaitis, P. V., . . . Widdicombe, J. (2007). ERS guidelines on the assessment of cough. *European Respiratory Journal*, 29(6), 1256-1276. doi: 10.1183/09031936.00101006
- Morice, A. H., Kastelik, J. A., & Thompson, R. (2001). Cough challenge in the assessment of cough reflex. *British Journal of Clinical Pharmacology*, 52(4), 365-375.
- Morice, A. H., McGarvey, L., & Pavord, I. (2006). Recommendations for the management of cough in adults. *Thorax*, 61 Suppl 1, i1-24. doi: 10.1136/thx.2006.065144

- Paintal, H. S., & Kushner, W. G. (2007). Aspiration syndromes: 10 clinical pearls every physician should know. *International Journal of Clinical Practice*, 61(5), 846-852. doi: 10.1111/j.1742-1241.2007.01300.x
- Pecova, R., Javorkova, N., Kudlicka, J., & Tatar, M. (2007). Tussigenic agents in the measurement of cough reflex sensitivity. *Journal of Physiology & Pharmacology*, 58 Suppl 5(2), 531-538.
- Pikus, L., Levine, M. S., Yang, Y. X., Rubesin, S. E., Katzka, D. A., Laufer, I., & Geftter, W. B. (2003). Videofluoroscopic studies of swallowing dysfunction and the relative risk of pneumonia. [Article]. *American Journal of Roentgenology*, 180(6), 1613-1616.
- Prudon, B., Birring, S. S., Vara, D. D., Hall, A. P., Thompson, J. P., & Pavord, I. D. (2005). Cough and glottic-stop reflex sensitivity in health and disease. *Chest*, 127(2), 550-557. doi: 10.1378/chest.127.2.550
- Puhl, R., & Brownell, K. D. (2001). Bias, Discrimination, and Obesity. *Obesity*, 9(12), 788-805.
- Ramsey, D., Smithard, D., & Kalra, L. (2005). Silent aspiration: What do we know? *Dysphagia*, 20(3), 218-225. doi: 10.1007/s00455-005-0018-9
- Romei, V., De Haas, B., Mok, R. M., & Driver, J. (2011). Auditory stimulus timing influences perceived duration of co-occurring visual stimuli. *Frontiers in Psychology*, 2. doi: 10.3389/fpsyg.2011.00215
- Schmid, G., & Ziegler, W. (2006). Audio-visual matching of speech and non-speech oral gestures in patients with aphasia and apraxia of speech. *Neuropsychologia*, 44(4), 546-555. doi: 10.1016/j.neuropsychologia.2005.07.002
- Scott, A., Perry, A., & Bench, J. (1998). A study of interrater reliability when using videofluoroscopy as an assessment of swallowing. *Dysphagia*, 13(4), 223-227. doi: 10.1007/pl00009576

- Shams, L. (2011). Early integration and bayesian causal inference in multisensory perception. In M. M. Murray (Ed.), *The neural bases of multisensory processes* (pp. 217-231). Hoboken: CRC Press.
- Shannon, R., Baekey, D. M., Morris, K. F., Nuding, S. C., Segers, L. S., & Lindsey, B. G. (2004). Production of reflex cough by brainstem respiratory networks. *Pulmonary Pharmacology & Therapeutics*, 17(6), 369-376. doi: 10.1016/j.pupt.2004.09.022
- Smith Hammond, C. A., & Goldstein, L. B. (2006). Cough and aspiration of food and liquids due to oral-pharyngeal dysphagia: ACCP evidence-based clinical practice guidelines. *Chest*, 129(1 Suppl), 154S-168S. doi: 10.1378/chest.129.1\_suppl.154S
- Smith Hammond, C. A., Goldstein, L. B., Horner, R. D., Ying, J., Gray, L., Gonzalez-Rothi, L., & Bolser, D. C. (2009). Predicting aspiration in patients with ischemic stroke: Comparison of clinical signs and aerodynamic measures of voluntary cough. *Chest*, 135(3), 769-777. doi: 10.1378/chest.08-1122
- Smith Hammond, C. A., Goldstein, L. B., Zajac, D. J., Gray, L., Davenport, P. W., & Bolser, D. C. (2001). Assessment of aspiration risk in stroke patients with quantification of voluntary cough. *Neurology*, 56(4), 502-506.
- Smith, J. A., Ashurst, H. L., Jack, S., Woodcock, A. A., & Earis, J. E. (2006). The description of cough sounds by healthcare professionals. *Cough*, 2, 1. doi: 10.1186/1745-9974-2-1
- Smith, J. A., Earis, J., & Woodcock, A. (2006). Establishing a gold standard for manual cough counting: video versus digital audio recordings. *Cough*, 2(1), 6. doi: 10.1186/1745-9974-2-6
- Wakasugi, Y., Tohara, H., Hattori, F., Motohashi, Y., Nakane, A., Goto, S., . . . Uematsu, H. (2008). Screening test for silent aspiration at the bedside. *Dysphagia*, 23(4), 364-370. doi: 10.1007/s00455-008-9150-7

- Widdicombe, J. G. (2008). A brief overview of mechanisms of cough. In K. F. Chung, J. G. Widdicombe & H. Boushey (Eds.), *Cough: Causes, mechanisms and therapy* (pp. 17-23): Wiley-Blackwell.
- Widdicombe, J. G., Addington, W. R., Fontana, G. A., & Stephens, R. E. (2011). Voluntary and reflex cough and the expiration reflex; implications for aspiration after stroke. *Pulmonary Pharmacology & Therapeutics*, 24(3), 312-317. doi: 10.1016/j.pupt.2011.01.015
- Widdicombe, J. G., & Fontana, G. A. (2006). Cough: what's in a name? *European Respiratory Journal*, 28(1), 10-15. doi: 10.1183/09031936.06.00096905
- Wright, C. E., Jackson, J., Thompson, R. L., & Morice, A. H. (2010). Validation of the ERS standard citric acid cough challenge in healthy adult volunteers. *Cough*, 6, 8. doi: 10.1186/1745-9974-6-8

## Appendix

### Appendix 1. Questionnaire

1. Age:

	Please ✓
Less than 25 years old	
26 – 35 years old	
36 – 50 years old	
51 – 75 years old	
76 years old and above	

2. To your knowledge, do you have normal hearing?

	Please ✓
Yes	
No	

3. To your knowledge, do you have normal vision or corrected-to-normal vision (wearing glasses)?

	Please ✓
Yes	
No	

4. In which setting do you work? (Please ✓)

<input type="checkbox"/>	Acute hospital	<input type="checkbox"/>	Rehab hospital
<input type="checkbox"/>	Chronic care	<input type="checkbox"/>	Outpatient clinic
<input type="checkbox"/>	Private practice	<input type="checkbox"/>	Home health care
<input type="checkbox"/>	School	<input type="checkbox"/>	Others



5. How many years have you been a practicing Speech Language Therapist/ Pathologist?

	Please ✓
Less than 1 year	
1 – 3 years	
4 – 6 years	
7 – 9 years	
10 – 12 years	
13 years +	

6. What type of population have you been managing? (In the past year)

	Please✓	Weight (%)		Weight (%)
Adult			Dysphagia	
Child				

7. How long have you been managing clients with dysphagia?

	Please ✓
Less than 1 year	
1 – 3 years	
4 – 6 years	
7 – 9 years	
10 – 12 years	
13 years +	

8. Do you listen to the client's cough when performing dysphagia assessment?

	Please ✓
Yes	
No	

9. Have you ever carried out cough reflex testing before?

	Please ✓	
Yes		If yes, proceed to question 10
No		If no, proceed to question 11

10. How long have you been carrying out cough reflex testing?

	Please ✓
Less than 1 year	
1 – 3 years	
4 – 6 years	
7 – 9 years	
10 – 12 years	
13 years +	

11. Have you ever had training in cough reflex testing?

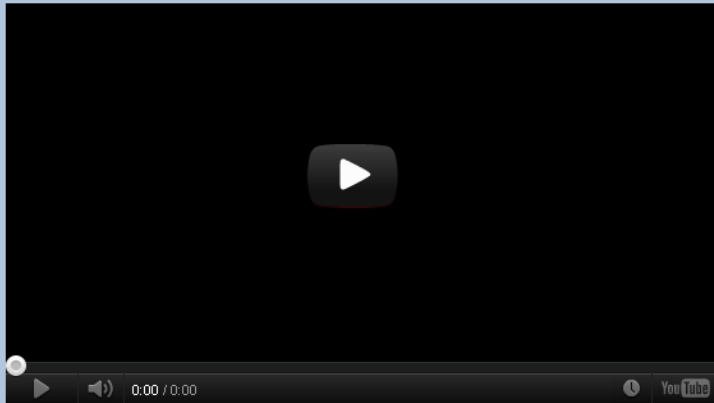
	Please ✓
Yes	
No	

12. In which country do you work?

.....

## Appendix 2. A view of the Online Site

Please see the clip and respond to the questions below



Is there an absence or presence of cough?

Absent

☐

Present

☐

If present, is it a strong or a weak cough?

Strong

☐

Weak

☐

Please comment on why you made this decision